Lab 1: Warming up Python with Historical Crypto

Please read through the whole lab before starting it.

Due: Sunday, September 30th, 2018 at 11:55pm

Group: This lab should be completed on your own. High level discussion is encouraged per the class collaboration policy.

Background. In this lab, you will explore cracking classical ciphers programmatically. This allows for a review of python to prepare you for upcoming labs. The three ciphers you will crack include the Caesar, Vigenère, and mono-alphabetic substitution cipher.

Environment. You should use python to implement any code for this lab. Along with the built-in Python functions you may **only** use the math and string libraries. Do not use any other libraries, online tools, cracking programs, or code you did not write.

Task In the tarred file encrypted.tar.gz you will find 9 files. Each of these files has been encrypted with a historic cipher. Your task is to provide a report with the following information for each file.

1) File name
2) Encryption key (not the decryption key)
    a) For Caesar this should be a number between 0 and 25 (assume A = 0, B = 1, C = 2, and so on)
    b) For Vigerene this should be a word
        i) I assumed that key words are one indexed (A+A = B, A+Z = A)
    c) For the mono alphabetic this should be a 26 character string
        i) The first letter should map from A
        ii) The second letter should map from B
        iii) If more than one letter is not used simply put a * for each unused letter
3) Decrypted text – spaces must be inserted
4) Process used to decrypt the file – this should be a simple description of the process you used to crack the cipher, referencing any code you wrote
5) Code you wrote to decrypt the file

Useful algorithms You may want to consider implementing the following algorithms to help with decrypting the files

1) A Caesar Cipher decryption algorithm

2) A Vigerene decryption algorithm
3) A mono-alphabetic substitution decryption algorithm
4) Letter frequency counter
5) Chi-squared test
6) Index of coincidence
7) Digraph frequency count
8) Repeat analysis to determine Vigerene key length

Helpful hints:

1) Some files have spaces removed
2) The input texts for files include English texts
3) Texts come from project Gutenberg and a couple other sources
4) At least one input text includes scientific names
5) Vigerene ciphers have key lengths of 5, 9, and 13
a. At least one is a proper noun
b. At least one is a sentence without spaces
6) Some English text does not follow normal frequency distribution

Submission. Include a report with the information for each file you decrypted. Make sure that you describe your cracking process in enough detail to demonstrate you understand how to crack the cipher. Also, make sure a classmate could reproduce your results from your process description and code.